

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Approximations of algebraic irrationalities with matrices

### This is the author's manuscript

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1700357> since 2020-04-14T09:33:59Z

*Published version:*

DOI:10.1080/10586458.2018.1526722

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Approximations of algebraic irrationalities with matrices

Stefano Barbero, Umberto Cerruti, Nadir Murru  
University of Turin, Department of Mathematics  
Via Carlo Alberto 10, Turin, Italy  
stefano.barbero@unito.it, umberto.cerruti@unito.it, nadir.murru@unito.it

## Abstract

We discuss the use of matrices for providing sequences of rationals that approximate algebraic irrationalities. In particular, we study the regular representation of algebraic extensions, proving that the ratios between two entries of the matrix of the regular representation converge to specific algebraic irrationalities. As an interesting special case, we focus on cubic irrationalities giving a generalization of the Khovanskii matrices for approximating cubic irrationalities. We discuss the quality of such approximations considering both rate of convergence and size of denominators. Moreover, we briefly perform a numerical comparison with well-known iterative methods (such as Newton and Halley ones), showing that the approximations provided by regular representations appear more accurate for the same size of the denominator.

*Keywords:* algebraic irrationals, diophantine approximation, matrices, root finding methods.

*AMS Subject Classification:* 11K60, 11J68.

## 1 Convergence properties for regular representations of algebraic extensions

The study of approximations of irrational numbers by means of rationals is a very important and rich research field. This research field is named *Diophantine approximation* in honor of Diophantus of Alexandria whose studies principally deal with researching rational solutions of algebraic equations. During the years, mathematicians have considerably improved results about Diophantine approximation.

In this context iterative methods, such as Newton method and higher order generalizations (i.e., Householder methods [8]) are widely used and studied. Recently, many different iterative root-finding methods have been developed improving classical methods (see, e.g., [1], [13], [7]). However, iterative methods are computationally slow and denominator size of the provided rational approximations rapidly increases. On the other hand, continued fractions provide best approximations of real numbers. However, their use is not ever convenient from a computational point of view.

In the case of algebraic numbers, iterative methods can be replaced by more convenient ones. For example, in [15], the authors propose an algorithm based on the LLL-reduction procedure for approximating algebraic numbers. Recently, different techniques involving powers of  $2 \times 2$  matrices have been developed for approximating quadratic irrationalities (see, e.g., [17] and [3]). In [9] and [11], authors introduced particular  $3 \times 3$  matrices for studying approximations of cubic roots. The use of matrices is very useful since power of matrices can be fastly evaluated and their entries are linear recurrent sequences whose properties can be exploited to study convergence. Moreover, study of simultaneous approximations is a very classical and well-studied topic, see, e.g., [4] and [6].

In the following, firstly, we introduce a family of matrices starting from the regular representation of algebraic extensions, studying their approximating properties. Then in section 2, we focus on cubic irrationalities, generalizing Khovanskii matrices and other kinds of matrices used in the approximation of cubic irrationalities. Moreover, In section 3, we provide numerical results about the studied approximations. In particular, we discuss performances of our approximations with respect to some parameters and we compare them with well-known iterative methods, such as Newton, Halley, and Noor methods.

Let  $\alpha$  be a real root of  $f(t) = t^m - \sum_{s=0}^{m-1} u_{m-s}t^s$ , with  $u_i \in \mathbb{Q}$ , for  $i = 1, \dots, m$ , irreducible over  $\mathbb{Q}$ . The algebraic extension  $\mathbb{Q}(\alpha)$  has basis  $(1, \alpha, \alpha^2, \dots, \alpha^{m-1})$ . Let  $\sum_{i=0}^{m-1} x_i \alpha^i$  be an element of  $\mathbb{Q}(\alpha)$ , it can be represented by the  $m \times m$  matrix  $M = (M_{ij})$  such that

$$\sum_{i=0}^{m-1} x_i \alpha^i \alpha^{j-1} = \sum_{i=1}^m M_{ij} \alpha^{i-1}, \quad j = 1, \dots, m. \quad (1)$$

The matrix  $M$  is usually called the regular representation of  $\mathbb{Q}(\alpha)$ . Let us observe that the above identities can be written also in the case that  $f(t)$  is reducible. Thus, in the following, we do not restrict  $f(t)$  to be irreducible and we formally define the matrix  $M$  by means of (1). Sometimes we will use the

notation  $M(\mathbf{x}, \mathbf{u})$ , where  $\mathbf{x} = (x_0, \dots, x_{m-1})$  and  $\mathbf{u} = (u_1, \dots, u_m)$ .

Let  $A$  be the companion matrix of  $f(t)$  defined as

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & u_m \\ 1 & 0 & 0 & \cdots & 0 & 0 & u_{m-1} \\ 0 & 1 & 0 & \cdots & 0 & 0 & u_{m-2} \\ \dots & \dots & \ddots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & u_3 \\ 0 & 0 & 0 & 0 & 1 & 0 & u_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & u_1 \end{bmatrix}$$

By definition of  $M$ , it follows that  $M = \sum_{n=0}^{m-1} x_n A^n$ . The entries of matrices  $A^n$  can be explicitly written, see, e.g., Theorem 3.1 in [5] (note that here the companion matrix is written in a slightly different form). In this way, we can explicitly write the entries of the matrix  $M$  as  $M_{ij} = \sum_{n=0}^{m-1} x_n a_{ij}^{(n)}$ , where

$$a_{ij}^{(n)} = \sum_{k_1+2k_2+\dots+mk_m=n-i+j} \frac{k_{m+1-i} + k_{m+2-i} + \dots + k_m}{k_1 + k_2 + \dots + k_m} \binom{k_1 + k_2 + \dots + k_m}{k_1, \dots, k_m} u_1^{k_1} u_2^{k_2} \dots u_m^{k_m}$$

for  $k_1, \dots, k_m$  non-negative integers and  $\binom{k_1 + k_2 + \dots + k_m}{k_1, \dots, k_m}$  is the multinomial coefficient.

In the following theorem we show convergence properties of  $M$  by means of the Vandermonde matrix.

**Theorem 1.** *Let  $\alpha_1, \dots, \alpha_m$  be distinct roots of  $f(t) = t^m - \sum_{s=0}^{m-1} u_{m-s} t^s$ . Let  $V = V(\alpha_1, \dots, \alpha_m)$  and  $M(\mathbf{x}, \mathbf{u})$  be the Vandermonde matrix of  $f(t)$  and the matrix defined by (1) with  $\alpha = \alpha_k$  for a given  $1 \leq k \leq m$ . Let us define*

$$c(\mathbf{x}, \alpha_k) = \min \left\{ \frac{|\sum_{i=0}^{m-1} x_i \alpha_k^i|}{|\sum_{i=0}^{m-1} x_i \alpha_j^i|} : j = 1, \dots, m, \alpha_j \neq \alpha_k \right\}.$$

*If  $c(\mathbf{x}, \alpha_k) > 1$ , then*

$$\lim_{n \rightarrow +\infty} \frac{M_{ij}^n}{M_{pq}^n} = \frac{V_{ik}^{-1} V_{kj}}{V_{pk}^{-1} V_{kq}},$$

*given any index  $i, j, p, q \in \{1, \dots, m\}$  (such that  $i \neq p$  and/or  $j \neq q$ ).*

*Proof.* Let  $V = V(\alpha_1, \dots, \alpha_m)$  be the Vandermonde matrix of  $f(t)$ , i.e.,

$$V = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{m-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{m-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \alpha_m & \alpha_m^2 & \cdots & \alpha_m^{m-1} \end{bmatrix}$$

It is well-known that  $V$  can be used in order to diagonalize the companion matrix of  $f(t)$  (see, e.g., [10] pag. 69), i.e.,

$$D(\alpha_1, \dots, \alpha_m) := \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_m \end{bmatrix} = VAV^{-1}.$$

Since  $M$  can be written as combination of the companion matrix, we have

$$VMV^{-1} = D\left(\sum_{i=0}^{m-1} x_i \alpha_1^i, \dots, \sum_{i=0}^{m-1} x_i \alpha_m^i\right)$$

and

$$VM^nV^{-1} = D\left(\left(\sum_{i=0}^{m-1} x_i \alpha_1^i\right)^n, \dots, \left(\sum_{i=0}^{m-1} x_i \alpha_m^i\right)^n\right).$$

From the previous identity, we obtain

$$M_{ij}^n = V_{i1}^{-1}V_{1j}\left(\sum_{i=0}^{m-1} x_i \alpha_1^i\right)^n + \dots + V_{im}^{-1}V_{mj}\left(\sum_{i=0}^{m-1} x_i \alpha_m^i\right)^n$$

and finally

$$\frac{M_{ij}^n}{M_{pq}^n} = \frac{V_{i1}^{-1}V_{1j}\left(\sum_{i=0}^{m-1} x_i \alpha_1^i\right)^n + \dots + V_{im}^{-1}V_{mj}\left(\sum_{i=0}^{m-1} x_i \alpha_m^i\right)^n}{V_{p1}^{-1}V_{1q}\left(\sum_{i=0}^{m-1} x_i \alpha_1^i\right)^n + \dots + V_{pm}^{-1}V_{mq}\left(\sum_{i=0}^{m-1} x_i \alpha_m^i\right)^n}$$

from which the thesis easily follows dividing numerator and denominator by  $\left(\sum_{i=0}^{m-1} x_i \alpha_k^i\right)^n$ .  $\square$

Remark 1. The condition  $c(\mathbf{x}, \alpha_k) > 1$  in the above theorem allows to specify which is the limit of the ratio  $\frac{M_{ij}^n}{M_{pq}^n}$ , given any index  $i, j, p, q \in \{1, \dots, m\}$ . In particular, it specifies which root of the polynomial  $f(t) = t^m - \sum_{s=0}^{m-1} u_{m-s} t^s$  can be approximated using the matrix  $M$ , since  $V_{ik}^{-1} V_{kj}$  and  $V_{pk}^{-1} V_{kq}$ , given any index  $i, j, p, q \in \{1, \dots, m\}$ , are all quantities that can be written involving only the root  $\alpha_k$  (and its powers) of  $f(t)$ , as direct calculations on the entries of the Vandermonde matrix show. In section 2, we will see some examples regarding cubic polynomials. In the case of totally real cubic and irreducible polynomials, we can see that there always exists an index  $k$  such that the condition  $c(\mathbf{x}, \alpha_k) > 1$  holds. Thus, when the quantity  $\mathbf{x}$  is fixed, we can check the values of  $c(\mathbf{x}, \alpha_k)$ , for any index  $k$ , in order to know which specific root of  $f(t)$  we approximate using  $M$  with the chosen value of  $\mathbf{x}$ . On the other hand, we can choose the values of  $\mathbf{x}$  so that  $c(\mathbf{x}, \alpha_k) > 1$  for a specific root  $\alpha_k$  that we would like to approximate.

We study the rate of convergence of the ratios of the entries of  $M^n$  in the next theorem, where we will use the following notation:

$$L = \frac{A_k}{B_k} = \frac{V_{ik}^{-1} V_{kj}}{V_{pk}^{-1} V_{kq}},$$

with  $V$  the Vandermonde matrix as defined in Theorem 1. Moreover, we will consider

$$\gamma_j = \sum_{i=0}^{m-1} x_i \alpha_j^i$$

and

$$c^{-1}(\mathbf{x}, \alpha_k) = \max \left\{ \frac{|\sum_{i=0}^{m-1} x_i \alpha_j^i|}{|\sum_{i=0}^{m-1} x_i \alpha_k^i|} : j = 1, \dots, m, \alpha_j \neq \alpha_k \right\} = \max \left\{ \frac{|\gamma_j|}{|\gamma_k|} : j = 1, \dots, m, \alpha_j \neq \alpha_k \right\},$$

where  $\gamma_j^n$ 's are the roots of the characteristic polynomial of  $M^n$ .

**Theorem 2.** *Let  $\alpha_1, \dots, \alpha_m$  be distinct roots of  $f(t) = t^m - \sum_{s=0}^{m-1} u_{m-s} t^s$ . Let  $V = V(\alpha_1, \dots, \alpha_m)$  and  $M(\mathbf{x}, \mathbf{u})$  be the Vandermonde matrix of  $f(t)$  and the matrix defined by (1) with  $\alpha = \alpha_k$  for a given  $1 \leq k \leq m$ . If  $c(\mathbf{x}, \alpha_k) > 1$  (i.e.  $0 < c^{-1}(\mathbf{x}, \alpha_k) < 1$ ) and  $A_l B_k - A_k B_l \neq 0$  (where  $l$  is the index such that  $c^{-1}(\mathbf{x}, \alpha_k) = \frac{|\gamma_l|}{|\gamma_k|}$ ), then the order of convergence of  $\left| \frac{M_{ij}^n}{M_{pq}^n} - L \right|$  is  $O((c^{-1}(\mathbf{x}, \alpha_k))^n)$ .*

*Proof.* From the previous theorem, we know that

$$\frac{M_{ij}^n}{M_{pq}^n} = \frac{\sum_{s=1}^m A_s \gamma_s^n}{\sum_{s=1}^m B_s \gamma_s^n} = \frac{A_k + \sum_{s=1, s \neq k}^m A_s \left(\frac{\gamma_s}{\gamma_k}\right)^n}{B_k + \sum_{s=1, s \neq k}^m B_s \left(\frac{\gamma_s}{\gamma_k}\right)^n}.$$

Let  $l$  be the index such that  $c^{-1}(\mathbf{x}, \alpha_k) = \frac{|\gamma_l|}{|\gamma_k|}$ , setting  $y = \frac{\gamma_l}{\gamma_k}$ , we have

$$\begin{aligned} \frac{M_{ij}^n}{M_{pq}^n} &= B_k^{-1} (A_k + A_l y^n + o(y^n)) \left(1 + \frac{B_l}{B_k} y^n + o(y^n)\right)^{-1} = B_k^{-1} \left(A_k - \frac{A_k B_l}{B_k} y^n + A_l y^n + o(y^n)\right) \\ &= \frac{A_k}{B_k} + \frac{A_l B_k - A_k B_l}{B_k^2} y^n + o(y^n). \end{aligned}$$

Hence, we have

$$\left| \frac{M_{ij}^n}{M_{pq}^n} - L \right| = \left| \frac{A_l B_k - A_k B_l}{B_k^2} y^n + o(y^n) \right| \leq \left| \frac{A_l B_k - A_k B_l}{B_k^2} \right| |y|^n + o(|y|^n) \leq C (c^{-1}(\mathbf{x}, \alpha_k))^n,$$

for a certain constant  $C$  depending on  $\left| \frac{A_l B_k - A_k B_l}{B_k^2} \right|$ . □

Remark 2. The above theorem holds under the condition that  $A_l B_k - A_k B_l \neq 0$ . We can observe that  $A_l B_k - A_k B_l = 0$  occurs in some special cases. In particular, there are some ratios of elements of  $M^n$  that are constant quantities. In fact, for any index  $t$ , we can see that

$$\frac{A_t}{B_t} = \frac{1}{(-1)^{m-1} \prod_{h=1}^m \alpha_h},$$

for  $i = m, j = m-1, p = 1, q = m$ . This follows from the following identities:

$$\begin{cases} V_{mt}^{-1} = \frac{1}{\prod_{\substack{1 \leq h \leq m \\ h \neq t}} (\alpha_t - \alpha_h)}, & V_{tm-1} = \alpha_t^{m-2} \\ V_{1t}^{-1} = (-1)^{m-1} \frac{\prod_{i=1, i \neq t}^m \alpha_i}{\prod_{\substack{1 \leq h \leq m \\ h \neq t}} (\alpha_t - \alpha_h)}, & V_{tm} = \alpha_t^{m-1} \end{cases}.$$

Thus, we are in a case where we have  $A_l B_k - A_k B_l = 0$ , but this corresponds to have  $\frac{M_{ij}^n}{M_{pq}^n} = \frac{M_{mm-1}^n}{M_{1m}^n}$  constant. Similarly, we can see that

$$\frac{A_t}{B_t} = \frac{1}{(-1)^{m-1} \prod_{h=1}^m \alpha_h},$$

for  $i = m, j = 1, p = 1, q = 2$  and any index  $t$ .

In the next section, we focus on the cubic case, since some well-known and studied matrices arise as particular cases of the matrix  $M$ .

## 2 Approximations of cubic irrationalities

The following matrix

$$\begin{pmatrix} x & r \\ 1 & x \end{pmatrix},$$

for  $x, r \in \mathbb{Z}$  and  $r$  positive square-free, is used to determine classic Rédei rational functions [16]. Powers of this matrix yield rational approximations of  $\sqrt{r}$ . In [3], the authors proved that among these approximations, Padé and Newton approximations can be found. A natural generalization of this matrix is given by

$$A = \begin{pmatrix} x & r & r \\ 1 & x & r \\ 1 & 1 & x \end{pmatrix}$$

for  $x, r \in \mathbb{Z}$  and  $r$  cube-free. This matrix has been introduced by Khovanskii [9] to approximate  $\sqrt[3]{r}$  and  $\sqrt[3]{r^2}$ . Let  $A_{ij}^n$  denote the  $ij$ -th entry of  $A^n$ , we have

$$\lim_{n \rightarrow +\infty} \frac{A_{11}^n}{A_{31}^n} = \sqrt[3]{r^2}, \quad \lim_{n \rightarrow +\infty} \frac{A_{21}^n}{A_{31}^n} = \sqrt[3]{r}.$$

In [11], authors studied the role of  $x$  in order to ensure the fastest convergence.

In [2], authors focused on

$$B = \begin{pmatrix} x & r & 0 \\ 0 & x & r \\ 1 & 0 & x \end{pmatrix}$$

and similarly, we have

$$\lim_{n \rightarrow +\infty} \frac{B_{11}^n}{B_{31}^n} = \sqrt[3]{r^2}, \quad \lim_{n \rightarrow +\infty} \frac{B_{21}^n}{B_{31}^n} = \sqrt[3]{r}.$$



Moreover, in this case the authors proved that  $\frac{B_{11}^n}{B_{31}^n}$  and  $\frac{B_{21}^n}{B_{31}^n}$  are convergents of certain generalized continued fractions yielding periodic representations of cubic roots. Finally, let  $\alpha$  be the real root largest in modulus of  $t^3 - pt^2 - qt - r$ , with  $p, q, r \in \mathbb{Q}$ . In [12], the author showed that matrix

$$C = \begin{pmatrix} x & r & pr \\ 0 & q+x & pq+r \\ 1 & p & p^2+q+x \end{pmatrix}$$

yields simultaneous rational approximations of  $\alpha - p$  and  $\frac{r}{\alpha}$ . However, in [12] the author did not focus on the study of rational approximations, but studied matrix  $C$  in order to determine periodic representations for any cubic irrational.

Matrices  $A$ ,  $B$ , and  $C$  are all particular cases of the matrix  $M$  studied in the previous section. Indeed, if we consider the cubic polynomial  $t^3 - pt^2 - qt - r$ ,  $p, q, r \in \mathbb{Q}$ , then, given integer numbers  $x, y, z$ , we have

$$M((x, y, z), (p, q, r)) = \begin{bmatrix} x & rz & ry + prz \\ y & x + qz & qy + (pq + r)z \\ z & y + pz & x + py + (p^2 + q)z \end{bmatrix}. \quad (2)$$

Previous matrices are particular cases of  $M$ . Indeed,

$$A = M((x, 1, 1), (0, 0, r)), \quad B = M((x, 0, 1), (0, 0, r)), \quad C = M((x, 0, 1), (p, q, r)).$$

By Theorem 1, it is possible to explicitly write limits of ratios between two elements of  $(M((x, y, z), (p, q, r)))^n$ . Let  $\alpha_1, \alpha_2, \alpha_3$  be roots of  $t^3 - pt^2 - qt - r$  and suppose  $c((x, y, z), \alpha_1) > 1$ , i.e., we have chosen  $x, y, z$  so that matrix  $M$  can be used for approximating  $\alpha_1$ . For instance, we have

$$\lim_{n \rightarrow +\infty} \frac{M_{22}^n}{M_{pq}^n} = \bar{M}_{pq},$$

given any  $p = 1, 2, 3$  and  $q = 1, 2, 3$ , where

$$\bar{M} = \begin{pmatrix} -\frac{\alpha_1(\alpha_2 + \alpha_3)}{\alpha_2\alpha_3} & -\frac{\alpha_2 + \alpha_3}{\alpha_2\alpha_3} & -\frac{\alpha_2 + \alpha_3}{\alpha_1\alpha_2\alpha_3} \\ \alpha_1 & 1 & \frac{1}{\alpha_1} \\ -\alpha_1(\alpha_2 + \alpha_3) & -\alpha_2 - \alpha_3 & -\frac{\alpha_2 + \alpha_3}{\alpha_1} \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1^2(\alpha_1 - p)}{r} & \frac{\alpha_1(\alpha_1 - p)}{r} & \frac{\alpha_1 - p}{r} \\ \alpha_1 & 1 & \frac{1}{\alpha_1} \\ \alpha_1(\alpha_1 - p) & \alpha_1 - p & \frac{\alpha_1 - p}{\alpha_1} \end{pmatrix}.$$

Another example is provided by

$$\lim_{n \rightarrow +\infty} \frac{M_{33}^n}{M_{pq}^n} = \bar{M}_{pq},$$

given any  $p = 1, 2, 3$  and  $q = 1, 2, 3$ , we have

$$\bar{M} = \begin{pmatrix} \frac{\alpha_1^2}{\alpha_2\alpha_3} & \frac{\alpha_1}{\alpha_2\alpha_3} & \frac{1}{\alpha_2\alpha_3} \\ -\frac{\alpha_1^2}{\alpha_2 + \alpha_3} & -\frac{\alpha_1}{\alpha_2 + \alpha_3} & -\frac{1}{\alpha_2 + \alpha_3} \\ \alpha_1^2 & \alpha_1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1^3}{r} & \frac{\alpha_1^2}{r} & \frac{\alpha_1}{r} \\ \frac{\alpha_1^2}{\alpha_1 - p} & \frac{\alpha_1}{\alpha_1 - p} & \frac{1}{\alpha_1 - p} \\ \alpha_1^2 & \alpha_1 & 1 \end{pmatrix}.$$

Clearly, if we have  $c((x, y, z), \alpha_2) > 1$ , then above results still hold exchanging indexes 1 and 2.

Remark 3. We can check the considerations of the Remark 2 in the above matrix. Indeed, we have that  $\frac{M_{31}}{M_{12}} = \frac{M_{32}}{M_{13}} = r$

### 3 Numerical results

In this section, we will deal with the quality of approximations provided by  $M$  comparing it with known iterative methods as Newton, Halley and similar ones. We would like to highlight that our method consists in evaluating powers of the matrix  $M$  and this is accomplished using only integer arithmetic, i.e., it is an error-free method.

It is well-known that continued fractions provide best approximations of real numbers (see [14] for a good survey about continued fractions). In particular, given the  $n$ -th convergent  $\frac{p_n}{q_n}$  of the continued fraction of a real number  $\alpha$ , then

$$\left| \alpha - \frac{p_n}{q_n} \right| \leq \left| \alpha - \frac{a}{b} \right|, \quad a, b \in \mathbb{Z},$$

for all  $b \leq q_n$ . However, evaluation of approximations by means of continued fractions is not generally an used method since a continued fraction is a non-terminating expression. Indeed, many different methods are studied and used in this context. In the particular case of approximations of algebraic numbers, many root-finding algorithms have been developed. Here, we

compare some of these methods with approximations provided by  $M$ . Taking into account classic definition of best approximations above described, we will compare quality of rational approximations, provided by different methods, having denominators with same size.

We will consider real roots of the Ramanujan cubic polynomial  $t^3+t^2-2t-1$  and we study approximations of  $M$  for different values of  $(x, y, z)$ . The roots of this polynomial are quite famous (see, e.g., [18]) and they are

$$\alpha_1 = 2 \cos \frac{2\pi}{7}, \quad \alpha_2 = 2 \cos \frac{4\pi}{7}, \quad \alpha_3 = 2 \cos \frac{8\pi}{7}$$

with  $|\alpha_3| > |\alpha_1| > |\alpha_2|$ .

### 3.1 Approximations of $M$ for different values of $(x, y, z)$

Considering  $M = M((x, y, z), (-1, 2, 1))$ , for some values of  $(x, y, z)$ , we provide approximations of  $\alpha_3$  by means of the sequence

$$m_n(x, y, z) := \frac{M_{21}^n(x, y, z)}{M_{31}^n(x, y, z)} - 1.$$

In Table 1 and 2, we summarize quality of our approximations for different values of  $(x, y, z)$ . In particular, we consider

$$\begin{cases} x = 0, y = 0, z = 1, & c(0, 0, 1) = 2.08815 \\ x = 1, y = -1, z = 1, & c(1, -1, 1) = 3.68141 \\ x = 0, y = -1, z = 1, & c(0, -1, 1) = 7.85086 \\ x = 69, y = 99, z = -124, & c(69, 99, -124) = 1343.4 \end{cases}$$

and we show values of  $|m_n(x, y, z) - \alpha_3|$  and size of denominators of  $m_n(x, y, z)$ , i.e., number of digits  $D_n(x, y, z)$  of  $M_{31}^n(x, y, z)$ .

In Figures 1 and 2, we depict the situations described in Tables 1 and 2, respectively.

We can observe that approximations provided by  $m_n(69, 99, -124)$  are the most accurate. However, they have the greatest denominators. Thus these approximations could not be optimal taking into account previous considerations about continued fractions.

It is interesting to observe that approximations  $m_n(0, -1, 1)$  are more accurate than  $m_n(1, -1, 1)$  and furthermore they have smaller denominators. Thus, approximations  $m_n(0, -1, 1)$  are surely better than approximations  $m_n(1, -1, 1)$  in any case.

In Table 3, we compare approximations whose denominators have the same number of digits (for 16, 35, and 62 digits). We can observe that approximations  $m_n(0, -1, 1)$  are more accurate than others with the same size of denominators.

In conclusion, if we want to obtain accurate approximations with low values of  $n$  and we are not interested in the size of denominators, it is sufficient to find values  $(x, y, z)$  that maximize  $c(x, y, z, \alpha_3)$ . However, in this way approximations could not be the better than others with the same size of denominators. Indeed, we have seen that approximations obtained in correspondence of  $c(0, -1, 1, \alpha_3) = 7.85086$  are more accurate than approximations with same size of denominators provided for  $c(69, 99, -124, \alpha_3) = 1343.4$ . It would be really interesting to study techniques that allow to determine values of  $(x, y, z)$  that provide best approximations in this sense.

### 3.2 Approximations of M for same values of $(x, y, z)$

By Theorem 1, we can obtain approximations of a cubic irrationality by using different ratios of  $M^n$ , for same values of  $(x, y, z)$ . Let us consider  $(x, y, z)$  such that  $c(x, y, z, \alpha_3) > 1$ , then the reader can check that

$$\lim_{n \rightarrow +\infty} \frac{M_{22}^n}{M_{21}^n} = \lim_{n \rightarrow +\infty} \frac{M_{23}^n}{M_{22}^n} = \lim_{n \rightarrow +\infty} \frac{M_{33}^n}{M_{32}^n} = \alpha_3.$$

In this paragraph, we briefly compare these approximations with each other. Let us consider  $(x, y, z) = (0, -1, 1)$ . In Tables 4 and 5, we report distance from exact value of  $\alpha_3$  and size of denominators for these approximations, respectively. We can see that there are not significative differences among these approximations. Similar results are obtained for other triples  $(x, y, z)$ .

### 3.3 Approximations of M for different values of $(p, q, r)$

By Theorem 1, we can also obtain approximations of a cubic irrationality using different values of  $(p, q, r)$ . In this paragraph, we focus on approximations of  $\alpha_2$ .

Considering  $(p, q, r) = (-1, 2, 1)$  we can find a triple  $(x, y, z)$  such that  $c(x, y, z, \alpha_2) > 1$ . For instance we have  $c(10, -2, -3, \alpha_2) = 2.67$  and we know that

$$\lim_{n \rightarrow +\infty} \frac{M_{21}^n}{M_{31}^n} - 1 = \alpha_2.$$

Moreover, we can consider polynomial  $t^3 + 2t^2 - t - 1$  (i.e., the reflected polynomial of  $t^3 + t^2 - 2t - 1$ ) whose roots are  $\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \frac{1}{\alpha_3}$ . In this case

we use  $(p, q, r) = (-2, 1, 1)$  and for  $(x, y, z) = (-3, 1, -1, 1/\alpha_2)$  we obtain  $c(x, y, z, 1/\alpha_2) = 2.67$ . Thus, by theorem 1, we have

$$\lim_{n \rightarrow +\infty} \frac{M_{11}^n}{M_{31}^n} - 1 = \alpha_2,$$

since  $r = 1$ . Let us observe that we have searched for triple  $(x, y, z)$  determining a value of  $c(x, y, z, 1/\alpha_2)$  similar to the previous case.

Finally, we can also consider  $(p, q, r) = (2, 1, -1)$ . In this case  $t^3 - 2t^2 - t + 1$  has roots  $\alpha_1 + 1$ ,  $\alpha_2 + 1$ ,  $\alpha_3 + 1$  and the reader can check that, e.g,  $c(x, y, z, \alpha_2 + 1) = 2.67$  so that

$$\lim_{n \rightarrow +\infty} \frac{M_{21}^n}{M_{31}^n} + 1 = \alpha_2,$$

since  $p = 2$ .

In Figures 3 and 4, we depict behavior of these approximations, considering differences with exact value of  $\alpha_2$  and size of denominators, respectively. Even in this case, there are not significative differences among these approximations. Thus, in general quality of approximations is heavily affected by values of  $c(x, y, z, \alpha)$ .

### 3.4 Comparison with known root-finding methods

In this paragraph, we compare approximations provided by  $M$  with Newton, Halley, and Noor [13] methods. We briefly recall these methods.

*Definition 1.* The Newton method provides rational approximations of a real root  $\alpha$  of  $f(t)$  by means of the sequence of rational numbers  $x_n$  by the equation

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad \forall n \geq 0$$

with a suitable initial condition  $x_0$ .

*Definition 2.* The Halley method provides rational approximations of a real root  $\alpha$  of  $f(t)$  by means of the sequence of rational numbers  $x_n$  by the equation

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{2(f'(x_n))^2 - f(x_n)f''(x_n)}, \quad \forall n \geq 0$$

with a suitable initial condition  $x_0$ .

*Definition 3.* The Noor method provides rational approximations of a real root  $\alpha$  of  $f(t)$  by means of the sequence of rational numbers  $x_n$  by the equations

$$\begin{cases} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, & \forall n \geq 0 \\ x_{n+1} = y_n - \frac{f(y_n)}{f'(y_n)} - \frac{(f(y_n))^2 f''(y_n)}{2(f'(y_n))^3}, & \forall n \geq 0 \end{cases}$$

with a suitable initial condition  $x_0$ .

In Table 6, we report approximations of  $\alpha_3$  provided by Newton, Halley, and Noor methods. In particular we report the size of denominators and the difference between these approximations and the exact value of the root.

Using notation of subsection 3.1, let us consider  $(p, q, r) = (-1, 2, 1)$  and  $(x, y, z) = (0, -1, 1)$ . Sequence  $m_n(x, y, z)$  approximates  $\alpha_3$ . In particular, we have that  $|m_{10} - \alpha_3| = 2.7 \times 10^{-9}$  and number of digits of denominator is 7. Furthermore, we have

$$|m_{25} - \alpha_3| = 1.0 \times 10^{-22}, \quad |m_{852} - \alpha_3| = 8.3 \times 10^{-763},$$

where denominators have 17 and 595 digits, respectively. Thus, approximations  $m_n$ , with same accuracy of iterative methods, have size of denominators much less than iterative methods. Equivalently, we can say that our approximations, having same size of denominators with respect to iterative methods, are much more accurate.

If we are only interested to have high accuracy in few steps, we can consider  $(p, q, r) = (-1, 2, 1)$ ,  $(x, y, z) = (69, 99, -124)$ ,  $N = M^3$  and  $N_n = \frac{N_{21}^n}{N_{31}^n} - 1$ . In Table 7, we report quality of approximations  $N_n$  for  $n = 1, \dots, 6$ . We can observe that in this case we reach high accuracy in few steps, with better performances than iterative methods.

Finally, we would like to observe that evaluation of powers of matrices is very fast from a computational point of view and it is faster than iterative methods.

## 4 Conclusion

We have introduced and studied a family of matrices (which generalize known ones) whose powers yield rational approximations of algebraic irrationalities. These matrices depend on some parameters whose meaning has been deeply discussed. These parameters allow to obtain many different approximations

for the same irrational, providing a very handy method that can be adjusted as necessary, in order to obtain the desired quality of approximation. Numerical results have been also presented in order to show effectiveness of our approach. Some questions should be deeper analyzed:

- study the role of  $\mathbf{x}$  in the size of denominators;
- explicitly determine maximum of  $c(\mathbf{x}, \alpha)$ ;
- study of the quality of simultaneous approximations (as defined, e.g., in [4]).

## 5 FIGURES AND TABLES

Table 1: Quality of approximations of matrices  $M$ :  $|m_n(x, y, z) - \alpha_3|$ .

	$n = 5$	$n = 20$	$n = 35$	$n = 50$	$n = 75$	$n = 100$
$ m_n(0, 0, 1) - \alpha_3 $	0.06	$9.8 \times 10^{-7}$	$1.6 \times 10^{-11}$	$2.5 \times 10^{-16}$	$2.5 \times 10^{-24}$	$2.6 \times 10^{-32}$
$ m_n(1, -1, 1) - \alpha_3 $	0.002	$1.2 \times 10^{-11}$	$3.8 \times 10^{-20}$	$1.2 \times 10^{-28}$	$8.7 \times 10^{-43}$	$6.1 \times 10^{-57}$
$ m_n(0, -1, 1) - \alpha_3 $	$8 \times 10^{-5}$	$3.1 \times 10^{-18}$	$1.2 \times 10^{-31}$	$4.4 \times 10^{-45}$	$1.9 \times 10^{-67}$	$7.9 \times 10^{-90}$
$ m_n(69, 99, -124) - \alpha_3 $	$1 \times 10^{-15}$	$4.0 \times 10^{-63}$	$9.5 \times 10^{-110}$	$8.6 \times 10^{-157}$	$6.1 \times 10^{-235}$	$3.7 \times 10^{-313}$

Table 2: Quality of approximations of matrices  $M$ : number of digits  $D_n(x, y, z)$  of  $M_{31}^n(x, y, z)$ .

Number of digits	$n = 5$	$n = 20$	$n = 35$	$n = 50$	$n = 75$	$n = 100$
$D_n(0, 0, 1)$	2	9	14	25	36	49
$D_n(1, -1, 1)$	4	16	21	39	59	78
$D_n(0, -1, 1)$	3	12	21	35	50	69
$D_n(69, 99, -124)$	13	52	92	135	203	269

Table 3: Comparison of approximations  $m_n(x, y, z)$  with same number of digits of denominators, for different values of  $(x, y, z)$ .

$(x, y, z)$	$n$	$ m_n(x, y, z) - \alpha_3 $	$D_n(x, y, z)$
$(0, 0, 1)$	37	$3.6 \times 10^{-12}$	16
$(1, -1, 1)$	20	$1.2 \times 10^{-11}$	16
$(0, -1, 1)$	23	$6.4 \times 10^{-21}$	16
$(69, 99, -124)$	6	$1.0 \times 10^{-19}$	16
$(0, 0, 1)$	74	$5.3 \times 10^{-24}$	35
$(1, -1, 1)$	45	$8.3 \times 10^{-26}$	35
$(0, -1, 1)$	50	$4.4 \times 10^{-45}$	35
$(69, 99, -124)$	14	$1.9 \times 10^{-44}$	35
$(0, 0, 1)$	128	$2.9 \times 10^{-41}$	62
$(1, -1, 1)$	82	$9.4 \times 10^{-47}$	62
$(0, -1, 1)$	91	$8.9 \times 10^{-82}$	62
$(69, 99, -124)$	23	$3.7 \times 10^{-72}$	62



Table 4: Distance from exact value of  $\alpha_3$  and different ratios of  $M$  with  $(x, y, z) = (0, -1, 1)$ .

	$n = 5$	$n = 20$	$n = 35$	$n = 50$	$n = 75$	$n = 100$
$ \frac{M_{21}^n}{M_{31}^n} - 1 - \alpha_3 $	$8.0 \times 10^{-5}$	$3.1 \times 10^{-18}$	$1.2 \times 10^{-31}$	$4.4 \times 10^{-45}$	$1.9 \times 10^{-67}$	$7.9 \times 10^{-90}$
$ \frac{M_{22}^n}{M_{21}^n} - \alpha_3 $	$5.0 \times 10^{-5}$	$2.1 \times 10^{-18}$	$8.1 \times 10^{-32}$	$3.0 \times 10^{-45}$	$1.3 \times 10^{-67}$	$5.4 \times 10^{-90}$
$ \frac{M_{23}^n}{M_{22}^n} - \alpha_3 $	$2.0 \times 10^{-5}$	$5.3 \times 10^{-19}$	$2.0 \times 10^{-32}$	$7.5 \times 10^{-46}$	$3.2 \times 10^{-68}$	$1.3 \times 10^{-90}$
$ \frac{M_{33}^n}{M_{32}^n} - \alpha_3 $	$2.1 \times 10^{-5}$	$7.6 \times 10^{-19}$	$2.9 \times 10^{-32}$	$1.1 \times 10^{-45}$	$4.6 \times 10^{-68}$	$1.9 \times 10^{-90}$

Table 5: Number of digits of  $M_{31}^n$ ,  $M_{21}^n$ ,  $M_{22}^n$ ,  $M_{32}^n$  with  $(x, y, z) = (0, -1, 1)$ .

Number of digits	$n = 5$	$n = 20$	$n = 35$	$n = 50$	$n = 75$	$n = 100$
$M_{31}^n$	3	12	21	35	50	69
$M_{21}^n$	3	14	25	35	50	70
$M_{22}^n$	4	14	25	35	53	70
$M_{32}^n$	4	14	25	34	53	70

Table 6: Quality of approximations of iterative methods.

Method	$n$	n. digits den.	$ x_n - \alpha_3 $
Newton	3	9	$1.1 \times 10^{-6}$
Newton	5	80	$9.2 \times 10^{-14}$
Newton	10	19352	$3.7 \times 10^{-762}$
Halley	2	9	$8.1 \times 10^{-8}$
Halley	3	45	$4.8 \times 10^{-22}$
Halley	6	28140	$1.2 \times 10^{-527}$
Noor	2	18	$1.1 \times 10^{-6}$
Noor	3	186	$2.7 \times 10^{-18}$
Noor	6	43136	$4.8 \times 10^{-471}$

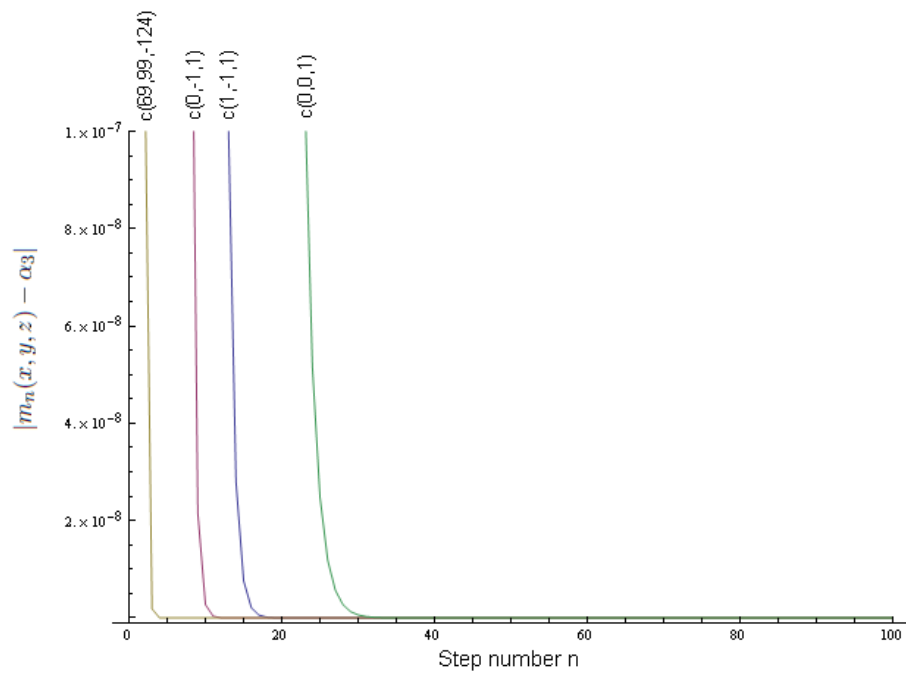


Figure 1: Rate of convergences of  $m_n(x, y, z)$  for different values of  $(x, y, z)$ .

Table 7: Quality of approximations of  $N_n$ .

$n$	n. digits den.	$ N_n - \alpha_3 $
1	8	$1.9 \times 10^{-9}$
2	24	$2.8 \times 10^{-28}$
3	73	$1.1 \times 10^{-84}$
4	219	$1.0 \times 10^{-253}$
5	658	$1.7 \times 10^{-760}$
6	1975	$8.4 \times 10^{-2281}$

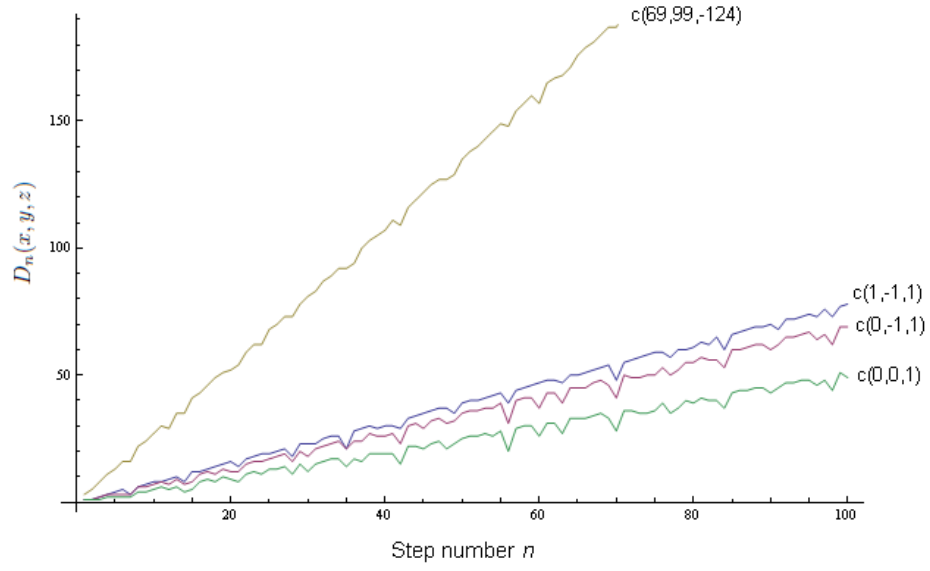


Figure 2: Number of digits of  $M_{31}^n$  for different values of  $(x, y, z)$ .

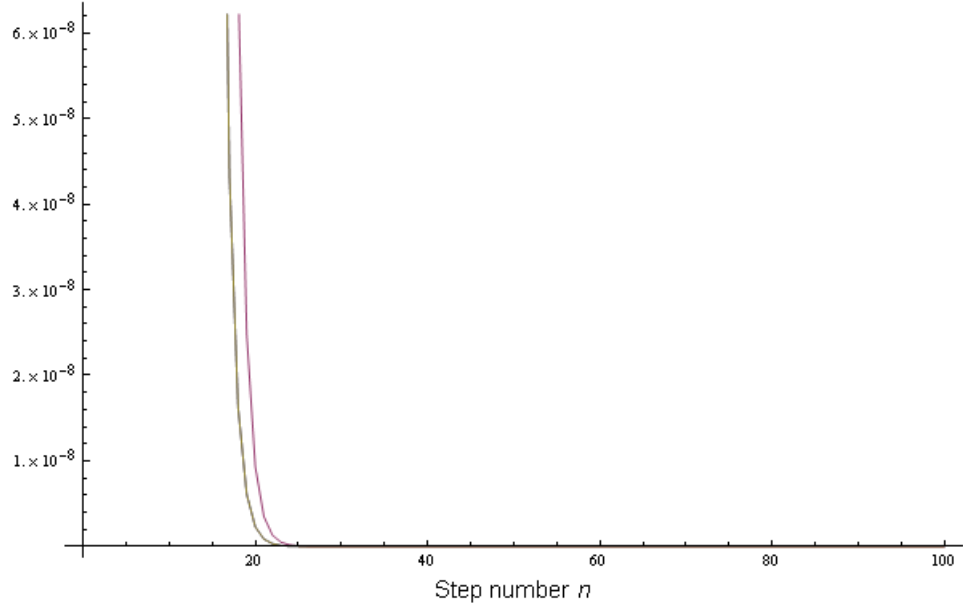


Figure 3: Differences between approximations provided by  $M$  and  $\alpha_2$  for different values of  $(p, q, r)$ :  $(-1, 2, 1)$ ,  $(-2, 1, 1)$ ,  $(2, 1, -1)$

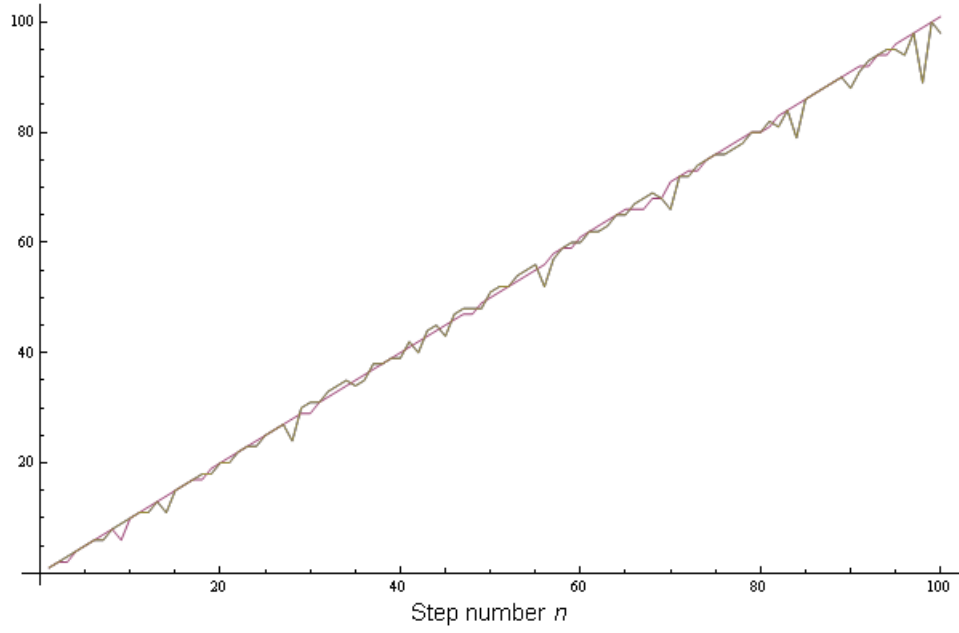


Figure 4: Number of digits of denominators for different values of  $(p, q, r)$ :  $(-1, 2, 1)$ ,  $(-2, 1, 1)$ ,  $(2, 1, -1)$

## References

- [1] Abbasbandy, S: Improving Newton–Raphson method for nonlinear equations by modified Adomian decomposition method, *Applied Mathematics and Computation*, Vol. 145, 887–893, (2003).
- [2] Abrate, M., Barbero, S., Cerruti, U., Murru, N.: Periodic representations for cubic irrationalities, *The Fibonacci Quarterly*, Vol. 50, No. 3, 252–264, (2012).
- [3] Abrate, M., Barbero, S., Cerruti, U., Murru, N.: Periodic representations and rational approximations of square roots, *Journal of Approximation Theory*, Vol. 175, 83–90, (2013).
- [4] Adams, W. W.: Simultaneous Diophantine approximations and cubic irrationals, *Pacific Journal of Mathematics*, Vol. 30, No. 1, 1–14, (1969).
- [5] W. Y. C. Chen, J. D. Louck, *The combinatorial power of the companion matrix*, Linear Algebra and its Applications, Vol. **232**, 261–278, 1996.
- [6] Chevallier, N.: Best simultaneous diophantine approximations of some cubic algebraic numbers, *Journal de théorie des nombres de Bordeaux*, Vol. 14, Issue 2, 403–414, (2002).
- [7] Grau, M., Diaz–Barrero, J. L.: An improvement to Ostrowski root–finding method, *Applied Mathematics and Computation*, Vol. 173, 450–456, (2006).
- [8] Householder, A. S.: *The Numerical Treatment of a Single Nonlinear Equation*, McGraw–Hill, New York, 1970.
- [9] Khovanskii, A. N.: *The Application of Continued Fractions and Their Generalizations to Problems in Approximation Theory*, cap. xii, 1963.
- [10] Lancaster, P., Tismenetsky, M.: *The theory of matrices*, Academic Press, 1985.
- [11] Mc Laughlin, J., Sury, B.: Some observations on Khovanskii’s matrix methods for extracting roots of polynomials, *Integers*, Vol. 7, Article A48, (2007).
- [12] Murru, N.: On the periodic writing of cubic irrationals and a generalization of Rédei functions, *International Journal of Number Theory*, Vol. 11, No. 11, 779–799, (2015).

- [13] Noor, K. I., Noor, M. A., Momani, S.: Modified Householder iterative method for nonlinear equations, *Applied Mathematics and Computation*, Vol. 190, 1534–1539, (2007).
- [14] Olds, C. D.: Continued fractions, Random House, 1963.
- [15] A. Pethő, M. E. Pohst, C. Bertók, On multidimensional Diophantine approximation of algebraic numbers, *Journal of Number Theory*, Vol. 171, 422–448, (2017).
- [16] Rédei, L.: Über eindeutige umkehrbare Polynome in endlichen Körper, *Acta Sci. Math. (Szeged)*, 11, 85–92, (1946).
- [17] Wildberger, N. J.: Pell’s equation without irrational numbers, *Journal of Integer Sequences*, Vol. 13, Article 10.4.3, (2010).
- [18] Witula, R., Slota, D.: New Ramanujan-type formulas and quasi-Fibonacci numbers of order 7, *Journal of Integer Sequences*, Vol. 10, Article 07.5.6, (2007).